AN APPLICATION OF MULTIPLE COMPARISON TECHNIQUES TO MODEL SELECTION

HIDETOSHI SHIMODAIRA

Department of Mathematical Engineering and Information Physics, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113, Japan

(Received April 15, 1996; revised March 26, 1997)

Abstract. Akaike's information criterion (AIC) is widely used to estimate the best model from a given candidate set of parameterized probabilistic models. In this paper, considering the sampling error of AIC, a set of good models is constructed rather than choosing a single model. This set is called a confidence set of models, which includes the minimum $\mathcal{E}\{\mathrm{AIC}\}$ model at an error rate smaller than the specified significance level. The result is given as P-value for each model, from which the confidence set is immediately obtained. A variant of Gupta's subset selection procedure is devised, in which a standardized difference of AIC is calculated for every pair of models. The critical constants are computed by the Monte-Carlo method, where the asymptotic normal approximation of AIC is used. The proposed method neither requires the full model nor assumes a hierarchical structure of models, and it has higher power than similar existing methods.

Key words and phrases: Akaike's information criterion, model selection, confidence set, multiple comparison with the best, Gupta's subset selection, variable selection, multiple regression, bootstrap resampling.